

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 30, 2002, 16:15:27 ; Search time 29 Seconds

(without alignments)  
461.830 Million cell updates/sec

Title: US-09-664-326-23

Perfect score: 368  
Sequence: 1 LHYPTCTESGNLCLCEGSN.....PKQSHNDGFEPRIPEYIQ 65

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21: \*  
1: sp\_archaea: \*  
2: sp\_bacteria: \*  
3: sp\_fungi: \*  
4: sp\_human: \*  
5: sp\_invertebrate: \*  
6: sp\_mammal: \*  
7: sp\_mhc: \*  
8: sp\_organelle: \*  
9: sp\_phage: \*  
10: sp\_plant: \*  
11: sp\_rodent: \*  
12: sp\_virus: \*  
13: sp\_vertebrate: \*  
14: sp\_unclassified: \*  
15: sp\_virus: \*  
16: sp\_bacteriophage: \*  
17: sp\_archaeop: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	247	67.1	84	5	007557	007557 hirudinaria
2	230	62.5	84	5	007558	007558 hirudinaria
3	138	37.5	25	5	09TWX5	09TWX5 hirudinaria
4	136	37.0	25	5	09TWX4	09TWX4 hirudinaria
5	74.5	20.2	1331	13	090W93	090W93 poecilia re
6	74	20.1	248	5	096962	096962 suberites d
7	74	20.1	2352	5	061240	061240 halocynthia
8	71	19.3	608	5	09VW80	09VW80 drosophila
9	70	19.0	169	2	08RP58	08RP58 anaplasma p
10	70	19.0	323	2	093NT8	093NT8 anaplasma p
11	69	18.8	315	13	09W600	09W600 gallus gall
12	68.5	18.6	2233	5	094711	094711 paramesicium
13	68	18.5	184	2	09RB30	09RB30 anaplasma p
14	68	18.5	278	2	09XCP8	09XCP8 anaplasma p
15	68	18.5	364	2	084968	084968 anaplasma p
16	66.5	18.1	516	3	0960Y3	0960Y3 saccharomyc

17	66.5	18.1	645	5	09VWR2	09VWR2 drosophila
18	66.5	18.1	762	13	042373	042373 brachydanio
19	66	17.9	761	4	060477	060477 homo sapien
20	66	17.9	1193	13	090819	090819 gallus gall
21	66	17.9	2150	5	044131	044131 caenorhabdl
22	65.5	17.8	213	5	09GCV8	09GCV8 plasmodium
23	65.5	17.8	248	5	096961	096961 geodia cydo
24	65.5	17.8	455	3	08X0R2	08X0R2 neospora
25	65.5	17.8	594	5	024970	024970 giardia lam
26	65.5	17.8	915	4	09BXN9	09BXN9 homo sapien
27	65.5	17.8	2601	4	09BX84	09BX84 homo sapien
28	65.5	17.8	2601	4	09BX85	09BX85 homo sapien
29	65.5	17.8	2624	4	09BX85	09BX85 homo sapien
30	65.5	17.8	2624	4	09BX87	09BX87 homo sapien
31	65.5	17.8	2689	4	09BX87	09BX87 homo sapien
32	65.5	17.8	2689	4	09BX85	09BX85 homo sapien
33	65.5	17.8	2721	4	09BX83	09BX83 homo sapien
34	65.5	17.8	2721	4	09BX90	09BX90 homo sapien
35	65.5	17.8	2724	4	09BX88	09BX88 homo sapien
36	65.5	17.8	2724	4	09BX89	09BX89 homo sapien
37	65.5	17.8	2812	4	09BX86	09BX86 homo sapien
38	65.5	17.8	2812	4	09BX88	09BX88 homo sapien
39	65	17.7	558	5	017797	017797 caenorhabdl
40	65	17.7	760	11	0925T8	0925T8 ratu
41	65	17.7	760	11	0920P3	0920P3 mus musculu
42	65	17.7	761	11	09QX10	09QX10 mus musculu
43	65	17.7	1043	5	017644	017644 caenorhabdl
44	65	17.7	1101	5	0964D2	0964D2 entamoeba h
45	65	17.7	1661	5	077244	077244 chlorohydra

## ALIGNMENTS

RESULT 1  
ID 007557 PRELIMINARY; PRT; 84 AA.  
AC 007557;  
DT 01-NOV-1998 (TREMURel. 08, Created)  
DT 01-NOV-1998 (TREMURel. 08, Last sequence update)  
DT 01-JUN-2001 (TREMURel. 17, Last annotation update)  
DE Hirudin HM2 precursor (Bufrudin).  
GN HM2.  
OS Hirudinaria manliensis (Buffalo leech).  
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinea; Hirudinidae; Hirudinea;  
OC Ancyrobellida; Hirudiniformes; Hirudinidae; Hirudinaria.  
OX NCBI\_TaxID=6419;  
[1]  
SEQUENCE FROM N.A., AND SEQUENCE OF 21-84.  
RP MEDLINE=93285156; PubMed=7685281;  
RX Sacherl E., Nittl G., Valsasina B., Orsini G., Visco C., Ferrera M.,  
RA Sawyer R.T., Sarmientos P.;  
RT "Novel hirudin variants from the leech Hirudinaria manliensis. Amino  
acid sequence, cDNA cloning and genomic organization.";  
RL Eur. J. Biochem. 214:295-304(1993).  
CC FUNCTION: HIRUDIN IS A POTENT THROMBIN-SPECIFIC PROTEASE  
INHIBITOR. IT FORMS A STABLE NON-COVALENT COMPLEX WITH ALPHA-  
THROMBIN, THEREBY ABOLISHING ITS ABILITY TO CLEAVE FIBRINOGEN.  
CC THROMBIN, BELONGS TO THE HIRUDIN FAMILY.  
CC -1- SIMILARITY: BELONGS TO THE HIRUDIN FAMILY.  
DR EMBL; X72786; CA51293.1; -;  
DR HSSP; P01050; SHIR.  
DR InterPro; IPR000429; Hirudin.  
DR Pfam; PF00713; Hirudin.1.  
DR PRINTS; PR00777; HIRUDIN.  
DR ProDom; PD004216; Hirudin.1.  
KW Serine protease inhibitor; Multigene family; Signal.  
FT SIGNAL 1 20  
FT CHAIN 21 84 HIRUDIN HM2.  
FT DISULFID 26 34 BY SIMILARITY.  
FT DISULFID 36 48 BY SIMILARITY.  
FT DISULFID 42 57 BY SIMILARITY.  
SQ SEQUENCE 84 AA; 9004 MW; CCBASDBSE7IBAF07 CRC64;



CC Echinayrote; Metacoa; Chordata; Craniata; Vertebrata; Neuteleostomi;  
CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neuteleostei;  
CC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;  
CC Cyprinodontiformes; Poeciliidae; Poecilia.  
OX NCBI\_taxid=8081;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=CAUDAL FIN;  
RA Ben J., Lim T.-M., Chan W.-K., Phang V.P.E.;  
RT "Molecular cloning of xanthine dehydrogenase gene from guppy (*Poecilia reticulata*).";  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY034103; AAK59659.1; -  
DR InterPro: IPR002888; 2fe-2s\_bind.  
DR InterPro: IPR000564; 2fe2s\_ferredoxin.  
DR InterPro: IPR000674; Aldxan\_dh.C.  
DR InterPro: IPR005107; CO\_deh\_flav\_C.  
DR InterPro: IPR002346; dehydrog\_moiyb.  
DR InterPro: IPR000572; Euk\_MD\_oxrd.  
DR InterPro: IPR010041; Ferredoxin.  
DR Pfam: PF01315; Ald\_xan\_dh.C; 1.  
DR Pfam: PF02738; Ald\_xan\_dh.C2; 1.  
DR Pfam: PF03450; CO\_deh\_flav\_C; 1.  
DR Pfam: PF00941; FAD\_binding\_5; 1.  
DR Pfam: PF00111; fer2; 1.  
DR Pfam: PF01799; fer2\_2; 1.  
DR ProDom: PD186071; 2fe-2s\_bind; 1.  
DR PROSITE: PS00197; 2FE2S\_FERREDOXIN; UNKNOWN\_1.  
DR PROSITE: PS00559; MOLYBDOPTEIN\_EUK; UNKNOWN\_1.  
KW Iron-sulfur.  
SQ SEQUENCE 1331 AA; 146336 MW; EBBE00AA6D36D126 CRC64;

Query Match	20.2%;	Score 74.5;	DB 13;	Length 1331;
Best Local Similarity	32.2%;	Pred. No. 2;		
Matches 19; Conservative	7;	Mismatches 22;	Indels 11;	Gaps 3

QY 2 TYTDTES-GQNLCLEG-----SNVGGQGNNKCLLGSDEKNQCVCGETPRQSH 51  
+ : : | | | | : | : : : |  
Db 138 TMDITQALGNLCRCCTGYRPIVDGCRTFCQEGNCC--qANGADCCCLNEGNTNESEH 194

RESULT 6	
096962	
ID 096962	PRELIMINARY;
000000	PRT; 248 AA.

DT	01-MAY-1999 (TREMblrel, 10, Created)
DT	01-MAY-1999 (TREMblrel, 10, last sequence update)
DT	01-MAR-2002 (TREMblrel, 20, last annotation update)

GN CD63R:  
OS Suberites domuncula (Sponge).  
OC Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;

```

OX NCBI_TaxID=55567;
RN [1]
RP SEQUENCE FROM N.A.

```

RT "Initiation of an aquaculture of sponges for their sustainable  
RT production of bioactive metabolites in open systems.";  
RL Mol. Mar. Biol. Biotechnol. 1:569-579(1999).  
DOI: 10.1007/s10226-000-0001-0

```
DR InterPro; IPR000301; Transmem_4.  
DR Pfam; PF00335; transmembrane4; 1.  
DR PRINTS; PR00259; TMFOUR.  
KW percent
```

Query Match	20.18;	Score 74;	DB 5;	length
Best Local Similarity	31.39;	Prod NC	0.28;	
SQ	SEQUENCE	248 AA;	26701 MW;	DB0F8D349A2E06EE CRC64;

	Matches	21, Conservative	10, Mismatches	18, Indels	18, Gaps	4, Ambiguous
QY	3	YTDCTESQNTL---	CICREGSNVCGGNCNICLISDSEKNCQVYGEQTPKPSHNDG---	DF	56	

Db 155 YTDYTDREHEEIPYSCICDDND--GDGAKCIPISBMGVNSTV-----YTDGCHDSF 202

QY 57 EEIPEEY 63  
: | | |

Db 203 VDFPREY 209

RESULT 7  
061240  
ID 061240 PRELIMINARY; PRT; 2352 AA

DT	01-AUG-1998	(TREMBlrel, 07, Created)
DT	01-AUG-1998	(TREMBlrel, 07, last sequence update)
DT	01-DEC-2001	(TREMBlrel, 19, last annotation update)

GN HRNOTCH.  
OS Halocynthia roretzi (Sea squirt).  
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;

```

OX      NCBI_TaxID=7729;
RN      [1]
RP      SEQUENCE FROM N.A.

```

RT "Notch homologue from *Halocynthia roretzi* is preferentially expressed  
RT in the central nervous system during ascidian embryogenesis.";  
RL Dev. Genes Evol. 207:371-380(1997).

DR HSSP; P00740; LEDM.  
DR InterPro; IPR002110; ANK.  
DR InterPro; IPR000152; Asx\_hydroxyl.

DR	InterPro; IPR000742; EGF_2.
DR	InterPro; IPR001881; EGF_Ca.
DR	InterPro; IPR001438; EGF_IT.

DR	Pfam; PF00023; ank; 6.
DR	Pfam; PF00008; EGF; 32.
DR	Pfam; PF00066; notch; 3.

DR PRINTS; PRO1452; NOTCH.  
DR SMART; SM00248; ANK; 5.  
DR SMART; SM00179; EGF\_CA; 17.

```
DR SMART; SM00004; NL; 2.
DR PROSITE; PS50088; ANK_REPEAT; 5.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
```

DR	PROSITE; PS00022; EGF_1; UNKNOWN_28.
DR	PROSITE; PS01186; EGF_2; 22.
DR	PROSITE; PS01187; EGF_CA; 18.

KW Hydroxylation; Repeat.  
SQ SEQUENCE 2352 AA; 252623 MW; 13DB1C056BB0D08D CRC64;

Query Match	20.1%	Score 74;	DB 5;	Length 2352;
Best Local Similarity	44.48%;	Pred. No. 4.3;		
Matches 20;	Conservative 4;	Mismatches 17;	Indels 4;	Gaps 3

**Qy**    2   TYTDCETSGQNLCCEGSNVCQGANKCILGSDGEKNQCVTG-EGT   45  
         | - | - | - | - | - | - : : - | - | - |  
**Db**    977 TGTHC-ESLON-LCTGPNIKNGGSCVQTSNTVSCNCILGYEET   1018

RESULT 8	
Q9VU80	
ID Q9VU80	PRELIMINARY;
AC Q9VU80	PRT; 608 AA

DT	01-MAY-2000 (Tremblure). 13, last sequence update)
DT	01-JUN-2001 (Tremblure). 17, last annotation update)
DE	CG10089 protein.
CU	CG10089

*Drosophila melanogaster* (Fruit fly).

```

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Gutter G.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.H.C., Blaise R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew K.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.D., Harvey D., Heiman T.J., Hernandez J.R., Houck C.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jajuli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paolel J.M.,
RA Palazuelo M., Peltman G.S., Pan S., Pollard J., Puri V., Reshe M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheibel F., Shen H.,
RA Shue B.C., Siden-Kimos I., Simpson M., Skupski M.P., Smith T.,
RA Spter E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003537; AAF49810.1; -.
DR HSSP: 016828; IMKP.
DR FlyBase: FBgn0036369; CG10089.
DR InterPro: IPR000340; DS_phosphatase.
DR InterPro: IPR000387; TYR_phosphatase.
DR Pfam: PF00782; DSPC; 1.
DR SMART: SM00195; DSPC; 1.
DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE: PS50054; TYR_PHOSPHATASE_DUAL; 1.
SQ SEQUENCE 608 AA; 6963 MW; 71EA135E5C7660D1 CRC64;

Query Match 19.3%; Score 71; DB 5; Length 608;
Best Local Similarity 31.8%; Pred. No. 2.3;
Matches 14; Conservative 10; Mismatches 18; Indels 2; Gaps 1;
OY 11 ONLCLCEGSGNVCQGNKCLGSDGKNCQVWGEGTPKPSHNDGDF-- 54
DB 199 GNRDICEGN--CSRGEKPTGADVDVGNVIGGEGEEDGEGEG 240

```

```

DE Major surface protein 2 hypervariable region (Fragment).
GN MS2.
OS Anaplasma phagocytophilum (Ehrlichia phagocytophila).
OC Bacteria; Proteobacteria; alpha subdivision; Rickettiales;
OC Anaplasmataceae; Anaplasma.
OX NCBI_TaxID=948;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BDS;
RX MEDLINE=21843110; PubMed=11854205;
RA Caspersen K., Park J.H., Patil S., Dunlter J.S.;
RT "Genetic variability and stability of Anaplasma phagocytophila msp2
RT (p4)."
RL Infect. Immun. 70:1230-1234(2002).
DR EMBL: AF443418; AAL77334.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 169 AA; 17639 MW; AB122A72A8F2CB70 CRC64;

Query Match 19.0%; Score 70; DB 2; Length 169;
Best Local Similarity 29.6%; Pred. No. 0.77;
Matches 21; Conservative 10; Mismatches 26; Indels 14; Gaps 3;
OY 6 CTESGONCLCEGSGNVCQGN-----KCLIGSGDEKN-QCVTGEGTPKPSHNDGDF-- 56
DB 64 GDNSSSTTTSGTNVSETGVFRDFIRATLKEDSKNWPSSGTPKPTNDNAKAVA 123
OY 57 -----EEIPPE 62
DB 124 KDVLQELTPEE 134

RESULT 10
O93NY8
ID O93NY8 PRELIMINARY; PRT; 323 AA.
AC O93NY8
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Major surface protein (Fragment).
OS Anaplasma phagocytophilum (Ehrlichia phagocytophila).
OC Bacteria; Proteobacteria; alpha subdivision; Rickettiales;
OC Anaplasmataceae; Anaplasma.
OX NCBI_TaxID=948;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WI 1;
RX MEDLINE=21320744; PubMed=11427556;
RA Lodes M.J., Mohamath R., Reynolds L.D., McNeill P., Kolbert C.P.,
RA Brulinsma E.S., Benson D.R., Hofmeister E., Reed S.G., Houghton R.L.,
RA Persing D.H.;
RT "Serodiagnosis of human granulocytic ehrlichiosis by using novel
RT combinations of immunoreactive recombinant proteins."
RL J. Clin. Microbiol. 39:2466-2476(2001).
DR EMBL: AF356509; AAK69699.1; -.
DR InterPro: IPR002566; Surface_Ag_msp4.
DR Pfam: PF01617; Surface_Ag_2; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 323 AA; 34533 MW; 51E0646ED4F8FDED CRC64;

Query Match 19.0%; Score 70; DB 2; Length 323;
Best Local Similarity 29.6%; Pred. No. 1.5;
Matches 21; Conservative 10; Mismatches 26; Indels 14; Gaps 3;
OY 6 CTESGONCLCEGSGNVCQGN-----KCLIGSGDEKN-QCVTGEGTPKPSHNDGDF-- 56
DB 185 GDNSSSTTTSGTNVSETGVFRDFIRATLKEDSKNWPSSGTPKPTNDNAKAVA 244
OY 57 -----EEIPPE 62
DB 245 KDVLQELTPEE 255

```

```

RESULT 11
ID 09600 PRELIMINARY; PRT; 315 AA.
AC 09600;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, last annotation update)
DE Flk protein.
GN FLK.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN 11
RP SEQUENCE FROM N.A.
RA Patel K.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE=96427388; PubMed=8812133;
RA Patel K., Connolly D., Anthor H., Nose K.;
RT "Cloning and early dorsal-axial expression of Flk, a chick
RT follistatin-related gene: Evidence for involvement in
RT dorso-lateral/neural induction. .";
RL Dev. Biol. 178:327-342(1996).
DR EMBL; AJ238977; CAB42968.1; -.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR002350; Kazal.
DR Pfam: PF00050; Kazal; 1.
DR SMART; SM00274; FOLD; 1.
DR SMART; SM00280; KAZAL; 1.
SQ SEQUENCE 315 AA; 35815 MW; C01889E005658A67 CRC64;

Query Match 18.8%; Score 69; DB 13; Length 315;
Best Local Similarity 33.3%; Pred. No. 2;
Matches 21; Conservative 6; Mismatches 26; Indels 10; Gaps 3;

QY 7 TESGONILCEGSGNVCQGN-----EGSNVCGGKNCILG-SPDEKNOCVTGEGTPKPSHNDGDFEEI 59
DB 45 TERGEPFCLCEGCKPHGRVYCGSNKTYLNHCEHLHDACITGS---KIVDYDGHCKEK 101
QY 60 PEE 62
DB 102 KSE 104

RESULT 12
ID 094711 PRELIMINARY; PRT; 2233 AA.
AC 094711;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, last annotation update)
DE 51C surface protein.
OS Parametium tetraurelia.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Pericollida;
OC Parametium.
OX NCBI_TaxID=5888;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=92106337; PubMed=1762150;
RA Nielsen E., You Y., Forney J.;
RT "Cysteine residue periodicity is a conserved structural feature of
RT variable surface proteins from Parametium tetraurelia. .";
RL J. Mol. Biol. 222:835-841(1991).
DR EMBL; M65164; AAA61740.1; -.
DR InterPro: IPR002895; Parametium_SA.
DR Pfam; PF01508; Parametium_SA; 25.
SQ SEQUENCE 2233 AA; 237078 MW; C064FE0AF7BB873B CRC64;

Query Match 18.6%; Score 68.5; DB 5; Length 2233;

```

```

Best Local Similarity 39.5%; Pred. No. 19;
Matches 17; Conservative 4; Mismatches 13; Indels 9; Gaps 3;

QY 6 CTESGONILCEGSGNVCQGN-----KCIIGSGENK-QCVTEGTPKPSHNDGDFEEI 45
DB 1830 CTIDGCG---CIGITCGKTNVNGCVTGIDGE---CITVAT 1866

RESULT 13
ID 09RB30 PRELIMINARY; PRT; 184 AA.
AC 09RB30;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, last annotation update)
DE 44 kDa major outer membrane protein (Fragment).
GN P44-15.
OS Anaplasma phagocytophilum (Ehrlichia phagocytophila).
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Anaplasmataceae; Anaplasma.
OX NCBI_TaxID=948;
RN 11
RP SEQUENCE FROM N.A.
RX STRAIN=H2;
RX MEDLINE=99292751; PubMed=10364227;
RA Zhi N., Ohashi N., Rikihisa Y.;
RT "Multiple p44 genes encoding major outer membrane proteins are
RT expressed in the human granulocytic ehrlichiosis agent. .";
RL J. Biol. Chem. 274:17828-17836(1999).
DR EMBL; AF135261; AAD41485.1; -.
DR InterPro: IPR002566; Surface_Ag_msp4.
DR Pfam; PF01617; Surface_Ag_2; 1.
FT NON_TER 1
FT NON_TER 184
SQ SEQUENCE 184 AA; 19245 MW; 2BC9AD4775FD55A1 CRC64;

Query Match 18.5%; Score 68; DB 2; Length 184;
Best Local Similarity 29.6%; Pred. No. 1.5;
Matches 21; Conservative 10; Mismatches 26; Indels 14; Gaps 3;

QY 6 CTESGONILCEGSGNVCQGN-----KCIIGSGENK-QCVTEGTPKPSHNDGDFEEI 56
DB 77 CGDNGSSSTITNSGANVSEGTGVFRDITRATLKEDGSKNWPSTSGTGPVYNDNAKAVA 136
QY 57 -----EEIPPEE 62
DB 137 KDLYQELTPPE 147

RESULT 14
ID 09XCP8 PRELIMINARY; PRT; 278 AA.
AC 09XCP8;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, last annotation update)
DE 44 kDa major outer membrane protein (Fragment).
GN P44-15.
OS Anaplasma phagocytophilum (Ehrlichia phagocytophila).
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Anaplasmataceae; Anaplasma.
OX NCBI_TaxID=948;
RN 11
RP SEQUENCE FROM N.A.
RX STRAIN=H2;
RX MEDLINE=99292751; PubMed=10364227;
RA Zhi N., Ohashi N., Rikihisa Y.;
RT "Multiple p44 genes encoding major outer membrane proteins are
RT expressed in the human granulocytic ehrlichiosis agent. .";
RL J. Biol. Chem. 274:17828-17836(1999).
DR EMBL; AF135261; AAD41480.1; -.
DR InterPro: IPR002566; Surface_Ag_msp4.
DR Pfam; PF01617; Surface_Ag_2; 1.

```

